

3D9

Page 1 of 7
RECEIVED
TECH CENTER 1600/2900
JUN 28 2002
1600

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/693,908

DATE: 01/26/2002

TIME: 13:36:55

Input Set : A:\ES.txt

Output Set: N:\CRF3\01262002\I693908.raw

ENTERED

3 <110> APPLICANT: HERMONAT, Paul L.
5 <120> TITLE OF INVENTION: ADENO-ASSOCIATED VIRUS AAV REP78 MAJOR REGULATORY PROTEIN,
MUTANTS
6 THEREOF AND USES THEREOF
8 <130> FILE REFERENCE: 023533/0130
10 <140> CURRENT APPLICATION NUMBER: US 09/693,908
11 <141> CURRENT FILING DATE: 2000-10-23
13 <150> PRIOR APPLICATION NUMBER: US 60/160,608
14 <151> PRIOR FILING DATE: 1999-10-21
16 <160> NUMBER OF SEQ ID NOS: 6
18 <170> SOFTWARE: PatentIn version 3.1
20 <210> SEQ ID NO: 1
21 <211> LENGTH: 26
22 <212> TYPE: DNA
23 <213> ORGANISM: Unknown
25 <220> FEATURE:
26 <223> OTHER INFORMATION: Primer
28 <400> SEQUENCE: 1
29 ccccgagggc cgaattcttt gtgcaa 26
32 <210> SEQ ID NO: 2
33 <211> LENGTH: 20
34 <212> TYPE: DNA
35 <213> ORGANISM: Unknown
37 <220> FEATURE:
38 <223> OTHER INFORMATION: Primer
40 <400> SEQUENCE: 2
41 acaagcagga ttgaaggcca 20
44 <210> SEQ ID NO: 3
45 <211> LENGTH: 20
46 <212> TYPE: DNA
47 <213> ORGANISM: Unknown
49 <220> FEATURE:
50 <223> OTHER INFORMATION: Primer
52 <400> SEQUENCE: 3
53 catatcacca gctcaccgtc 20
56 <210> SEQ ID NO: 4
57 <211> LENGTH: 110
58 <212> TYPE: DNA
59 <213> ORGANISM: Human papillomavirus
61 <400> SEQUENCE: 4
62 actacaataa ttcattgtata aaactaaggg cgtaaccgaa atcggttgaa ccgaaaccgg 60
64 ttagtataaa agcagacatt ttatgcacca aaagagaact gcaatgttc 110
67 <210> SEQ ID NO: 5
68 <211> LENGTH: 4675

RECEIVED

JUN 28 2002

TECH CENTER 1600/2900

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/693,908

DATE: 01/26/2002

TIME: 13:36:55

Input Set : A:\ES.txt

Output Set: N:\CRF3\01262002\I693908.raw

```

69 <212> TYPE: DNA
70 <213> ORGANISM: adeno-associated virus 2
72 <220> FEATURE:
73 <221> NAME/KEY: CDS
74 <222> LOCATION: (321)..(2183)
75 <223> OTHER INFORMATION:
78 <400> SEQUENCE: 5
79 ttggccactc cctctctgcg cgctcgctcg ctcaactgagg ccgggcgacc aaaggtcgcc      60
81 cgacgcccg gctttgcccg ggcggcctca gtgagcgagc gagcgcgag agagggagtg      120
83 gccaaactcca tcaactagggg ttcttgagg ggtggagtcg tgacgtgaat tacgtcatag      180
85 ggttagggag gtcctgtatt agaggtcacg tgagtgtttt gcgacatttt gcgacaccat      240
87 gtggtcacgc tgggtattta agcccagagtg agcacgcagg gtctccattt tgaagcggga      300
89 ggtttgaacg cgcagccgcc atg ccg ggg ttt tac gag att gtg att aag gtc      353
90 Met Pro Gly Phe Tyr Glu Ile Val Ile Lys Val
91 1 5 10
93 ccc agc gac ctt gac ggg cat ctg ccc ggc att tct gac agc ttt gtg      401
94 Pro Ser Asp Leu Asp Gly His Leu Pro Gly Ile Ser Asp Ser Phe Val
95 15 20 25
97 aac tgg gtg gcc gag aag gaa tgg gag ttg ccg cca gat tct gac atg      449
98 Asn Trp Val Ala Glu Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met
99 30 35 40
101 gat ctg aat ctg att gag cag gca ccc ctg acc gtg gcc gag aag ctg      497
102 Asp Leu Asn Leu Ile Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu
103 45 50 55
105 cag cgc gac ttt ctg acg gaa tgg cgc cgt gtg agt aag gcc ccg gag      545
106 Gln Arg Asp Phe Leu Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu
107 60 65 70 75
109 gcc ctt ttc ttt gtg caa ttt gag aag gga gag agc tac ttc cac atg      593
110 Ala Leu Phe Phe Val Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met
111 80 85 90
113 cac gtg ctc gtg gaa acc acc ggg gtg aaa tcc atg gtt ttg gga cgt      641
114 His Val Leu Val Glu Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg
115 95 100 105
117 ttc ctg agt cag att cgc gaa aaa ctg att cag aga att tac cgc ggg      689
118 Phe Leu Ser Gln Ile Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly
119 110 115 120
121 atc gag ccg act ttg cca aac tgg ttc gcg gtc aca aag acc aga aat      737
122 Ile Glu Pro Thr Leu Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn
123 125 130 135
125 ggc gcc gga ggc ggg aac aag gtg gtg gat gag tgc tac atc ccc aat      785
126 Gly Ala Gly Gly Gly Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn
127 140 145 150 155
129 tac ttg ctc ccc aaa acc cag cct gag ctc cag tgg gcg tgg act aat      833
130 Tyr Leu Leu Pro Lys Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn
131 160 165 170
133 atg gaa cag tat tta agc gcc tgt ttg aat ctc acg gag cgt aaa cgg      881
134 Met Glu Gln Tyr Leu Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg
135 175 180 185
137 ttg gtg gcg cag cat ctg acg cac gtg tcg cag acg gag gag cag aac      929

```

RAW SEQUENCE LISTING

DATE: 01/26/2002

PATENT APPLICATION: US/09/693,908

TIME: 13:36:55

Input Set : A:\ES.txt

Output Set: N:\CRF3\01262002\I693908.raw

```

279 ctcaacttca agctctttaa cattcaagtc aaagaggta cgcagaatga cggtagcagc 3193
281 acgattgcca ataaccttac cagcacggtt cagggtgtta ctgactcgga gtaccagctc 3253
283 ccgtacgtcc tcggctcggc gcatcaagga tgccctccgc cgttcccagc agacgtcttc 3313
285 atggtgccac agtatggata cctcaccctg aacaacggga gtcaggcagt aggacgtctc 3373
287 tcattttact gcctggagta ctttcttctc cagatgctgc gtaccggaaa caactttacc 3433
289 ttcagctaca cttttgagga cgttcttttc cacagcagct acgtcacag ccagagtctg 3493
291 gaccgtctca tgaatcctct catcgaccag tacctgtatt acttgagcag aacaaacact 3553
293 ccaagtggaa ccaccacgca gtcaaggctt cagttttctc aggccggagc gagtgcacatt 3613
295 cgggaccagt ctaggaactg gcttcttgga ccctgttacc gccagcagcg agtatcaaag 3673
297 acatctgcgg ataacaaca cagtgaatac tcgtggactg gagctaccaa gtaccacctc 3733
299 aatggcagag actctctggt gaatccggcc atggcaagcc acaaggacga tgaagaaaag 3793
301 tttttctc agagcggggt tctcatcttt gggaagcaag gctcagagaa aacaaatgtg 3853
303 aacattgaaa aggtcatgat tacagacgaa gaggaaatcg gaacaacca tcccgtggct 3913
305 acggagcagt atggttctgt atctaccaac ctccagagag gcaacagaca agcagctacc 3973
307 gcagatgtca acacacaagg cgttcttcca ggcattggtt ggcaggacag agatgtgtac 4033
309 cttcaggggc ccatctgggc aaagattcca cacacggagc gacattttca cccctctccc 4093
311 ctcatgggtg gattcggact taaacacct cctccacaga ttctcatcaa gaacaccccg 4153
313 gtacctgcga atccttcgac caccttcagt gcggcaaagt ttgcttctc catcacacag 4213
315 tactccacgg gacacggtca gcgtggagat cgagtgggag ctgcagaagg aaaacagcaa 4273
317 acgttggaat cccgaaattc agtacacttc caactacaac aagtctgtta atcgtggact 4333
319 tacgtggagt actaatggcg tgtattcaga gcctgcctcc attggcacca gatacctgac 4393
321 tcgtaactcg taattgcttg ttaatcaata aaccgtttaa ttcgtttcag ttgaactttg 4453
323 gtctctgcgt atttctttct tatctagttt ccatggctac gtagataagt agcatggcgg 4513
325 gttaatcatt aactacaagg aacccttagt gatggagttg gccactccct ctctgcgcgc 4573
327 tcgctcgtc actgaggccg ggcgacaaa ggctgcctcg cgcctgggct ttgcccgggc 4633
329 ggctcagtg agcgagcgag cgcgcagaga gggagtggcc aa 4675

```

332 <210> SEQ ID NO: 6

333 <211> LENGTH: 621

334 <212> TYPE: PRT

335 <213> ORGANISM: adeno-associated virus 2

337 <400> SEQUENCE: 6

```

339 Met Pro Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp
340 1 5 10 15
343 Gly His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu
344 20 25 30
347 Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile
348 35 40 45
351 Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu
352 50 55 60
355 Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val
356 65 70 75 80
359 Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu
360 85 90 95
363 Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile
364 100 105 110
367 Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu
368 115 120 125
371 Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly
372 130 135 140

```

RAW SEQUENCE LISTING

DATE: 01/26/2002

PATENT APPLICATION: US/09/693,908

TIME: 13:36:55

Input Set : A:\ES.txt

Output Set: N:\CRF3\01262002\I693908.raw

```

203      445      450      455
205 aag gtc acc aag cag gaa gtc aaa gac ttt ttc cgg tgg gca aag gat      1745
206 Lys Val Thr Lys Gln Glu Val Lys Asp Phe Arg Trp Ala Lys Asp
207 460      465      470      475
209 cac gtg gtt gag gtg gag cat gaa ttc tac gtc aaa aag ggt gga gcc      1793
210 His Val Val Glu Val Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala
211      480      485      490
213 aag aaa aga ccc gcc ccc agt gac gca gat ata agt gag ccc aaa cgg      1841
214 Lys Lys Arg Pro Ala Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg
215      495      500      505
217 gtg cgc gag tca gtt gcg cag cca tcg acg tca gac gcg gaa gct tcg      1889
218 Val Arg Glu Ser Val Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser
219      510      515      520
221 atc aac tac gca gac agg tac caa aac aaa tgt tct cgt cac gtg ggc      1937
222 Ile Asn Tyr Ala Asp Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly
223      525      530      535
225 atg aat ctg atg ctg ttt ccc tgc aga caa tgc gag aga atg aat cag      1985
226 Met Asn Leu Met Leu Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln
227 540      545      550      555
229 aat tca aat atc tgc ttc act cac gga cag aaa gac tgt tta gag tgc      2033
230 Asn Ser Asn Ile Cys Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys
231      560      565      570
233 ttt ccc gtg tca gaa tct caa ccc gtt tct gtc gtc aaa aag gcg tat      2081
234 Phe Pro Val Ser Glu Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr
235      575      580      585
237 cag aaa ctg tgc tac att cat cat atc atg gga aag gtg cca gac gct      2129
238 Gln Lys Leu Cys Tyr Ile His His Ile Met Gly Lys Val Pro Asp Ala
239      590      595      600
241 tgc act gcc tgc gat ctg gtc aat gtg gat ttg gat gac tgc atc ttt      2177
242 Cys Thr Ala Cys Asp Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe
243      605      610      615
245 gaa caa taaatgat ttt aaatcaggta tggctgccga tggttatctt ccagattggc      2233
246 Glu Gln
247 620
249 tcgaggacac tctctctgaa ggaataagac agtgggtggaa gctcaaacct ggcccaccac      2293
251 caccaaagcc cgcagagcgg cataaggacg acagcagggg tcttgtgctt cctgggtaca      2353
253 agtacctcgg acccttcaac ggactcgaca agggagagcc ggtcaacgag gcagacgccg      2413
255 cggccctcga gcacgtacaa agcctacgac cggcagctcg acagcggaga caaccgtac      2473
257 ctcaagtaca accacgccga cgcggagt ttt caggagcgcc ttaaagaaga tacgtctttt      2533
259 gggggcaacc tcggacgagc agtcttcag gcgaaaaaga gggttcttga acctctgggc      2593
261 ctggttgagg aacctgttaa gacggctccg ggaaaaaaga ggccggtaga gcactctcct      2653
263 gtggagccag actcctcctc gggaaccgga aaggcgggcc agcagcctgc aagaaaaaga      2713
265 ttgaattttg gtcagactgg agacgcagac tcagtacctg acccccagcc tctcggacag      2773
267 ccaccagcag cccctctcgg tctgggaact aatacgatgg ctacaggcag tggcgacca      2833
269 atggcagaca ataacgaggg cgcgcagcga gtgggtaatt cctccggaaa ttggcattgc      2893
271 gattccacat ggatgggcga cagagtcac accaccagca cccgaacctg ggccctgcc      2953
273 acctacaaca accacctcta caaacaatt tccagccaat caggagcctc gaacgacaat      3013
275 cactactttg gctacagcac cccttggggg tattttgact tcaacagatt cactgccac      3073
277 ttttcaccac gtgactggca aagactcatc aacaacaact ggggattccg acccaagaga      3133

```

RAW SEQUENCE LISTING

DATE: 01/26/2002

PATENT APPLICATION: US/09/693,908

TIME: 13:36:55

Input Set : A:\ES.txt

Output Set: N:\CRF3\01262002\I693908.raw

138	Leu	Val	Ala	Gln	His	Leu	Thr	His	Val	Ser	Gln	Thr	Gln	Glu	Gln	Asn	
139			190					195					200				
141	aaa	gag	aat	cag	aat	ccc	aat	tct	gat	gcg	ccg	gtg	atc	aga	tca	aaa	977
142	Lys	Glu	Asn	Gln	Asn	Pro	Asn	Ser	Asp	Ala	Pro	Val	Ile	Arg	Ser	Lys	
143		205					210					215					
145	act	tca	gcc	agg	tac	atg	gag	ctg	gtc	ggg	tgg	ctc	gtg	gac	aag	ggg	1025
146	Thr	Ser	Ala	Arg	Tyr	Met	Glu	Leu	Val	Gly	Trp	Leu	Val	Asp	Lys	Gly	
147	220					225					230					235	
149	att	acc	tcg	gag	aag	cag	tgg	atc	cag	gag	gac	cag	gcc	tca	tac	atc	1073
150	Ile	Thr	Ser	Glu	Lys	Gln	Trp	Ile	Gln	Glu	Asp	Gln	Ala	Ser	Tyr	Ile	
151					240					245					250		
153	tcc	ttc	aat	gcg	gcc	tcc	aac	tcg	cgg	tcc	caa	atc	aag	gct	gcc	ttg	1121
154	Ser	Phe	Asn	Ala	Ala	Ser	Asn	Ser	Arg	Ser	Gln	Ile	Lys	Ala	Ala	Leu	
155				255					260					265			
157	gac	aat	gcg	gga	aag	att	atg	agc	ctg	act	aaa	acc	gcc	ccc	gac	tac	1169
158	Asp	Asn	Ala	Gly	Lys	Ile	Met	Ser	Leu	Thr	Lys	Thr	Ala	Pro	Asp	Tyr	
159		270					275						280				
161	ctg	gtg	ggc	cag	cag	ccc	gtg	gag	gac	att	tcc	agc	aat	cgg	att	tat	1217
162	Leu	Val	Gly	Gln	Gln	Pro	Val	Glu	Asp	Ile	Ser	Ser	Asn	Arg	Ile	Tyr	
163		285					290					295					
165	aaa	att	ttg	gaa	cta	aac	ggg	tac	gat	ccc	caa	tat	gcg	gct	tcc	gtc	1265
166	Lys	Ile	Leu	Glu	Leu	Asn	Gly	Tyr	Asp	Pro	Gln	Tyr	Ala	Ala	Ser	Val	
167	300					305					310					315	
169	ttt	ctg	gga	tgg	gcc	acg	aaa	aag	ttc	ggc	aag	agg	aac	acc	atc	tgg	1313
170	Phe	Leu	Gly	Trp	Ala	Thr	Lys	Lys	Phe	Gly	Lys	Arg	Asn	Thr	Ile	Trp	
171					320					325					330		
173	ctg	ttt	ggg	cct	gca	act	acc	ggg	aag	acc	aac	atc	gcg	gag	gcc	ata	1361
174	Leu	Phe	Gly	Pro	Ala	Thr	Thr	Gly	Lys	Thr	Asn	Ile	Ala	Glu	Ala	Ile	
175				335					340					345			
177	gcc	cac	act	gtg	ccc	ttc	tac	ggg	tgc	gta	aac	tgg	acc	aat	gag	aac	1409
178	Ala	His	Thr	Val	Pro	Phe	Tyr	Gly	Cys	Val	Asn	Trp	Thr	Asn	Glu	Asn	
179			350					355					360				
181	ttt	ccc	ttc	aac	gac	tgt	gtc	gac	aag	atg	gtg	atc	tgg	tgg	gag	gag	1457
182	Phe	Pro	Phe	Asn	Asp	Cys	Val	Asp	Lys	Met	Val	Ile	Trp	Trp	Glu	Glu	
183		365					370					375					
185	ggg	aag	atg	acc	gcc	aag	gtc	gtg	gag	tcg	gcc	aaa	gcc	att	ctc	gga	1505
186	Gly	Lys	Met	Thr	Ala	Lys	Val	Val	Glu	Ser	Ala	Lys	Ala	Ile	Leu	Gly	
187	380					385					390					395	
189	gga	agc	aag	gtg	cgc	gtg	gac	cag	aaa	tgc	aag	tcc	tcg	gcc	cag	ata	1553
190	Gly	Ser	Lys	Val	Arg	Val	Asp	Gln	Lys	Cys	Lys	Ser	Ser	Ala	Gln	Ile	
191				400						405					410		
193	gac	ccg	act	ccc	gtg	atc	gtc	acc	tcc	aac	acc	aac	atg	tgc	gcc	gtg	1601
194	Asp	Pro	Thr	Pro	Val	Ile	Val	Thr	Ser	Asn	Thr	Asn	Met	Cys	Ala	Val	
195				415					420					425			
197	att	gac	ggg	aac	tca	acg	acc	ttc	gaa	cac	cag	cag	ccg	ttg	caa	gac	1649
198	Ile	Asp	Gly	Asn	Ser	Thr	Thr	Phe	Glu	His	Gln	Gln	Pro	Leu	Gln	Asp	
199			430					435					440				
201	cgg	atg	ttc	aaa	ttt	gaa	ctc	acc	cgc	cgt	ctg	gat	cat	gac	ttt	ggg	1697
202	Arg	Met	Phe	Lys	Phe	Glu	Leu	Thr	Arg	Arg	Leu	Asp	His	Asp	Phe	Gly	

```
=> s adeno?
  16 FILES SEARCHED...
L1    1759336 ADENO?

=> s l1 and vir?
  10 FILES SEARCHED...
  21 FILES SEARCHED...
  31 FILES SEARCHED...
L2    200183 L1 AND VIR?

=> s l2 and rep?
   6 FILES SEARCHED...
  16 FILES SEARCHED...
  22 FILES SEARCHED...
  26 FILES SEARCHED...
  27 FILES SEARCHED...
L3    92098 L2 AND REP?

=> s l3 and 78
L4    9548 L3 AND 78

=> s l4 and mutant
  32 FILES SEARCHED...
L5    4090 L4 AND MUTANT
```